RESEARCH ARTICLE

Performance of detection dogs and visual searches for scat detection and discrimination amongst related species with identical diets

Annegret Grimm-Seyfarth[1,2], Aleksandra Zarzycka[1,2,3], Teresa Nitz[1], Lisa Heynig[4], Nadine Weissheimer[5], Simone Lampa[1,6], Reinhard Klenke[1]

1 UFZ – Helmholtz Centre for Environmental Research, Department of Conservation Biology, Permoserstr. 15, 04318 Leipzig, Germany 2 Wildlife Detection Dogs e.V., Reichweinstraße 27, 74867 Neunkirchen, Germany 3 University of Warsaw, Faculty of Biology, Żwirki i Wigury 101, 02-089 Warsaw, Poland 4 Johanniter-Unfall-Hilfe e.V. Regionalverband Leipzig/Nordsachsen, Torgauer Str. 233, 04347 Leipzig, Germany 5 Im Herzen Hund, Arnimstr. 29/4, 23566 Lübeck, Germany 6 Wildtierökologie & Monitoring, Rübezahlweg 6, 04277 Leipzig, Germany

Corresponding author: Annegret Grimm-Seyfarth (annegret.grimm@ufz.de)

Abstract

Ecology often faces the problem that many threatened species are highly elusive but also conflict-laden. Thus, proper monitoring data are inevitable for their conservation and management. Indirect monitoring through scats is frequently used for such species, but scats of related species or species with similar diet are often visually indistinguishable. Since genetic methods for species identification are time-consuming and cost-intensive, a verification of the target species beforehand would be extremely beneficial in reducing effort to the analysis of the target species only. Such species discrimination could be provided through species-specific scat detection dogs. Therefore, we evaluated the reliability of species-specific scat detection dogs for two mustelid species feeding on identical diets: the Eurasian otter (*Lutra lutra*) and the American mink (*Neovison vison*), both of which are conflict-laden and increasing their populations and distribution ranges in central Europe. Their scats resemble each other in morphology and odour, exacerbating the differentiation even for experts. To evaluate whether detection dogs can reliably discriminate between related species feeding on similar diets and if their use would be beneficial, we tested their abilities against those of humans.

We first proved that scat characteristics are not statistically different between species. Likewise, visual species identification through people with different experience levels was only partly successful. Experts showed higher average accuracy (0.89) than non-experts (0.72 and below), but detection dogs (4 dogs) were able to discriminate otter and mink scats under laboratory conditions with an accuracy of 0.95. Moreover, otter scat detection dogs found up to four times more scat samples in the field, were twice as fast as human searchers and found an almost equal number of scats with different characteristics, while humans mostly found older and larger scats placed on hotspots. We conclude that using detection dogs for species identity will allow subsequent laboratory analyses to be species-specific and avoid spending time and money on laboratory work of the wrong species. It also provides more precise and unbiased information about the target species.

Introduction

Many animal and plant species of interest in nature conservation are hard to find and therefore difficult to monitor. This is especially true for elusive or nocturnal species as well as for other species that are hard to capture. Combined with the fact that many elusive species are threatened or endangered and conflict-laden at the same time (Henle et al. 2013), it is even more challenging to analyse these species' distributions or population statuses, securing their conservation and management. Therefore, species monitoring is often realised indirectly through visual scat monitoring (MacKay et al. 2008). However, scat samples of related species or species with similar diets often cannot be differentiated morphologically (Davison et al. 2002, MacKay et al. 2008). This implies that monitoring, relying on scat only, could overestimate the presence of the target species unless determined genetically. Since genetic methods are time-consuming and cost-intensive, a verification of the target species beforehand would be extremely beneficial in reducing effort to the analysis of the target species only.

Species discrimination can be provided using specially trained scat detection dogs, which are increasingly used for a reliable monitoring in nature conservation (see MacKay et al. 2008, Dahlgren et al. 2012, Long and MacKay 2012, Woollett (Smith) et al. 2014). In previous studies, dogs successfully distinguished sympatric, closely related species [e.g. black bear (*Ursus americanus*) from grizzly (*U. arctos*) (Hurt et al. 2000, Wasser et al. 2004); kit fox (*Vulpes macrotis mutica*) from red fox (*V. vulpes*) and coyote (*Canis latrans*) (Smith et al. 2003); and bobcat (*Lynx rufus*) from red fox, grey fox (*Urocyon cinereoargenteus*), kit fox, domestic dog (*C. familiaris*) and coyote (Harrison 2006)]. While the diets of the abovementioned sympatric species partially overlap, to our knowledge, dogs have never been systematically tested to discriminate between related species feeding on an almost identical diet.

One such example is the syntopic occurrence of the threatened Eurasian otter (*Lutra lutra*) and the invasive American mink (*Neovison vison*) in Central Europe. A substantial proportion of both species' diet can consist of commercial fish (Geidezis 1996, Jedrze-

jewska et al. 2001, Poledníková et al. 2013), making them highly conflict-laden species (Kranz 2000, Kruuk 2006, Klenke et al. 2013). Since their populations are increasing in Central Europe (Kranz 2000, Hauer et al. 2009, Klenke et al. 2013, Poledníková et al. 2013, Lampa et al. 2015), the potential for human-wildlife-conflicts is increasing as well and proper species monitoring is essential (Poledníková et al. 2013). For otters, the standard monitoring method is non-invasive population monitoring with scats (Mason and Macdonald 1987). However, scat samples of otter and mink are often similar in shape, colour, and odour. In fact, most mink scats, visually identified as otter scats, contained fish remains and did not smell like typical mink scats, such that even experts could not differentiate between both (Lampa et al. 2015). Genetic capture-recapture methods have already been performed throughout Europe (e.g. Arrendal et al. 2007, Hájková et al. 2009, Bonesi et al. 2013, Lampa et al. 2015). For example, between 2006 and 2008, only 3–5% of all potential otter scats collected in the Upper Lusatia, Germany, could be conclusively genetically assigned to syntopic mink individuals and, from 2010 to 2012, this number has increased up to 20% (Lampa et al. 2015).

Given the discrimination abilities of trained detection dogs, we evaluated the reliability and benefit of species-specific detection dogs in species monitoring using the Eurasian otter and American mink. Specifically, we systematically tested the accuracy and precision of detection dogs and humans for discrimination between related species, even when individuals were fed the same diet. Secondly, we compared relative and absolute detection abilities of dogs and humans under field conditions.

Materials and methods

Collection and comparison of Eurasian otter and American mink scat

To compare otter scat and otter-like mink scat (herein referred to as "mink scat"), we used the data from a previous otter monitoring study of the Upper Lusatian Heath and Pond Landscape in eastern Saxony, Germany (see Lampa et al. 2015 and Suppl. material 1: S1.1 for a description of the study site and sampling design). To investigate whether otter and mink scats differed in their characteristics (amount, colour, sliminess, exposure), we used Pearson's chi-squared test. If the test was significant, we used pairwise chi-squared tests and the Bonferroni-Holm correction for p-values as a post-hoc test (Holm 1979).

Additionally, we obtained scats from captive otter and mink from the otter centre Hankensbüttel (http://otterzentrum.de). In April 2015, eight Eurasian otters (4 males, 4 females) and four mink (all males) were fed the same fish species only for two weeks. After three days of digestion adjustment, three scat samples were collected per species per day for ten days, labelled according to the potential individuals and frozen immediately. A total of 60 frozen scat samples divided into three to five partial samples each were stored in separate, airtight plastic tubes. To determine the difference in scat colour between both species when fed on an identical diet, we performed Fisher's exact rank test using 20 random otter and mink samples.

Detection dogs and training

Dogs used in this study were privately-owned pet dogs. We trained two dogs to detect otter scats only. For comparisons, we trained two additional dogs to detect mink scats only (Table 1). We used positive reinforcement without coercion. While three dogs were already plant detection or rescue dogs, the youngest one had no previous experience. All dogs first performed lab tests (see below) on a scent box (Fig. 1). The scent box, constructed of a cardboard, had five holes with a centre-to-centre distance of 30 cm to ensure scent overlap (McKay 2014). Five plastic cups could be placed into the scent box deep enough that the handler could not see whether the cup contained anything. Additionally, the cups were equipped with a lid with three small holes. We used the systematically collected scats from captivity for training and testing on the scent box to ensure that individuals were fed on an identical diet.

All dogs were trained on scat scent following Wasser et al. (2004). For each training day, scat samples needed were defrosted, used for training and subsequently refrozen. For imprinting, i.e. focussing a dog on its target scent, we used mixed samples of four individuals of the target species (Kerley 2004). We chose a passive alert (sitting) that involves no barking, scratching or pawing to keep the influence on targets and wildlife in later field deployments as low as possible (DeMatteo et al. 2019). For applications in field-work, we also added a "show-me" command where the dog pointed to the specific scat to be able to detect hidden scats and species-specific scats on hotspots (defined as minimum three scats in less than 15 cm distance).

In contrast to Wasser et al. (2004), the dogs were not guided along the scent box, but allowed to walk along independently for three times at a maximum upon a "search" command (Kerley 2004) (Fig. 1). Each correct finding was quickly rewarded with food. Dogs were allowed to play with their toy following several trials before they got a break. After the dogs were habituated to the scent box and trained on their target scat, we performed a pre-test to examine that the dog had understood his task (Suppl. material 1: S1.2). When the dog passed that test without an error, species discrimination training started. We used the method of errorless discrimination (Gadbois and Reeve 2014) to train the dogs to reliably detect target scent and ignore non-target scent, i.e. otter scat detection dogs alert at otter but ignore mink scat, while the mink scat detection dogs alert at mink but ignore otter scat (Suppl. material 1: S1.2). Then, the final discrimination test followed (see below).

Once the detection and discrimination were working properly indoors, we started the field training of the otter scat detection dogs. We did not train the mink scat detection dogs in the field since the overall goal was to improve necessary otter monitoring. Searching took place without a leash such that the dog would not be biased towards a particular direction by the handler (MacKay et al. 2008, Woollett (Smith) et al. 2014). At the beginning of field training, the handler knew locations where the target and non-target scats were hidden so that the dog could be rewarded immediately. Subsequent trials were conducted blind as illustrated in Fig. 2.

Throughout field training, the dogs were confronted with various surfaces (short or tall grass, understorey, debris, logs, concrete) and elevations (level to steep) and trained at different temperatures (-5 °C to 32 °C), precipitation (sunny, cloudy, drizzle, rain-

Table 1. Overview of the scat detection dogs used in this study. Ages refer to the tests in 2017 and not to the start of the training.

Dog	Bagheera	Zammy	Cue	Zoey
Age	8 years	1 year	8 years	4 years
Breed	Australian Cattle Dog Mix	Border Collie	Border Collie	Border Collie
Origin	animal shelter	breeder	breeder	breeder
Previous experience	plant detection dog	none	certified rescue dog	certified rescue dog
Target species	Eurasian otter	Eurasian otter	American mink	American mink
Handler	AZ	AGS	LH	LH

Figure 1. Searching along the scent box (left) and alerting at the target scat (right), taking the example of the otter scat detection dog Zammy.

ing, snowing) and wind conditions (none to severe). To enlarge the dogs' scent range, we used genetically verified scat samples of male and female otters from the Upper Lusatia additionally to the samples from captivity. After both dogs were able to detect all otter and ignore all mink scats in ten subsequent blind trials, whereby each trial contained 1–2 otter and 1–2 mink scats, we also took the dogs to areas where Eurasian otters but no American minks were present to mimic field conditions and verify whether the dogs would alert on wild scats. We obtained permission from pond farmers and local game authorities for all pond areas where we conducted field training.

Species discrimination tests

We assessed human and dog abilities to discriminate between otter and mink scats using 20 random captive scat samples of different otter and mink individuals all fed with fish only. By chance, eleven samples belonged to otter and nine belonged to mink. The 20 samples were unknown to both humans and dogs. Every test person and every dog got the same samples for comparability. Neither test persons nor handlers knew how many samples belonged to which species.

Test persons were classified according to their expertise into beginner (person has never worked with scats of otters or minks or just started, i.e. total hours in the field: 0–72), experienced (person has already searched for at least one species, i.e. total hours in the field: ca. 240–336) and expert (person has already searched for both species intensively,
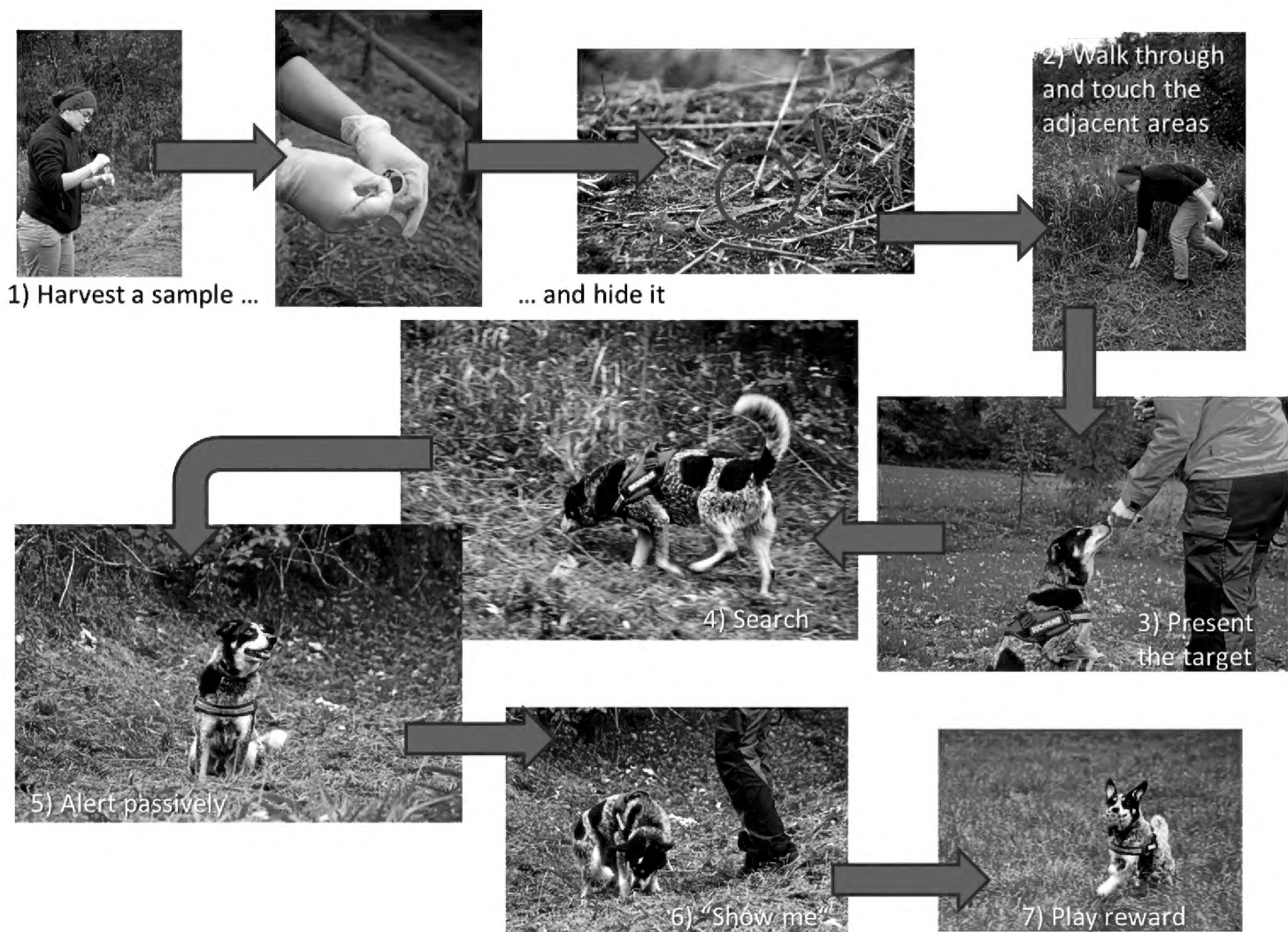
1) Harvest a sample …   … and hide it   2) Walk through and touch the adjacent areas

4) Search   3) Present the target

5) Alert passively   6) "Show me"   7) Play reward

Figure 2. Chronology in the field training of the scat detection dogs, taking the example of the otter scat detection dog Bagheera. (1) A field assistant hid the scat using laboratory gloves and (2) walked through the search area a few times in order to cover up tracks out of sight of the handler. (3) Ritualised, the dog sat in front of the handler and the target scent was presented to the dog in a tube. (4) The dog was sent to search the area independently and (5) gave a passive alert (sitting or lying down, depending on the dog's preferences) in front of the scat found. (6) We requested the command "show me" where the dog indicated precisely the source of scent through freeze-and-stare but was carefully trained not to touch the scat. The assistant would then indicate the handler whether this find was correct. (7) For each correctly found scat, the dog would be rewarded with a valuable treat (e.g. dried meat or favourite toy). Wrong alerts would be commented with "wrong" and the search would start again.

i.e. total hours in the field > 840). Test samples were presented in plastic cups in a 4 × 5 grid (Fig. 3) and test persons were asked to assign each scat sample to either otter or mink. To obtain which parameters influenced correct species assignment of a scat, we performed a generalised linear mixed model (GLMM) with "1" implying correct and "0" incorrect assignment using a binomial error distribution. Explanatory parameters were tested for non-collinearity and include the species, sex, the day sampled reflecting the time since digestion adjustment, colour and mass of the scat, expertise and the random intercept of the person's ID. We obtained parameter significances by means of a likelihood ratio test (LRT) of the full model against the model without the parameter in question and overall model significance using an LRT of the full model against a model including the random term only. Since humans assigned the species mainly visually, we further analysed whether test candidates used the scat colour for species assignment (Suppl. material 1: S1.3).
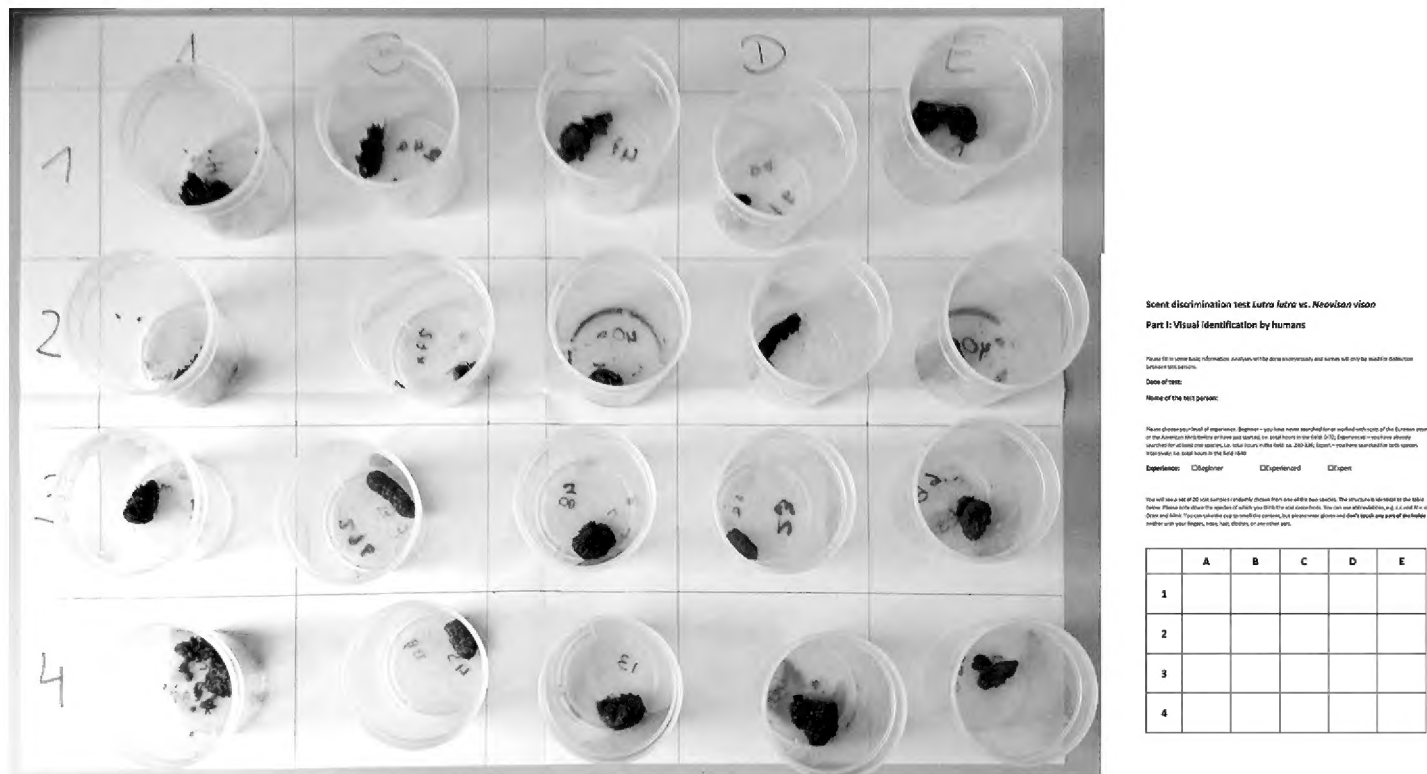
Figure 3. Setup for human scent discrimination test.

To ensure a standardised testing procedure between dog and human abilities, we used the scent box for final discrimination tests with dogs (Fig. 1). We created test protocols with 20 random trials. Each trial contained one cup with an individual otter scat and one cup with an individual mink scat as well as three blanks (Suppl. material 1: S1.2). We ensured that each of the 20 test samples was used at least once. The target and non-target species were correctly assigned if the dog alerted to or ignored that cup, respectively. All tests were conducted double-blind (Suppl. material 1: S1.2).

For both human and dog species discrimination tests, we calculated the sensitivity as the amount of true-positives (target samples correctly assigned) divided by the total number of all target samples and the specificity as the amount of true-negatives (non-target samples correctly assigned) divided by the total number of non-target samples. We further calculated the accuracy (validity) as the amount of all true-positives and true-negatives divided by the total number of target and non-target samples presented and the precision (reliability) as the amount of true-positives divided by the total number of positive responses (all alerts for dogs or all otter assignments for humans). We refer to a target sample as otter scat for humans and otter detection dogs, but mink scat for mink detection dogs and a non-target sample as mink scat for humans and otter detection dogs, but otter scat for mink detection dogs.

Field procedures in the Upper Lusatia

In March 2017, we conducted a five-day otter monitoring in the Upper Lusatia in the study area of Lampa et al. (2015) where both otter and mink occur (see Suppl. material 1: S1.1). Our overall aim was to compare visual and detection dog searches for wild otter

scat under field conditions. Since both cleared and uncleared transect monitoring is regularly used for otter monitoring, we applied both methods. We thus divided the monitoring season in one day for clearance of four pre-defined transects where all scats found visually were removed, two days to compare the abilities to detect freshly deposited wild otter scat from the night before on the pre-defined transects amongst teams and two days for otter presence monitoring on uncleared transects at ponds nearby. Transects followed trails along ponds in four pond areas: "Biwatsch", "Entenschenke", "Mühlteich", and "Langner". They were characterised by a starting point and direction (Suppl. material 1: S1.1, Fig. S1). Vegetation was comparable amongst transects and consisted mainly of low grass and trees. "Biwatsch" also contained a substantial amount of bushes.

We allowed two hours for clearance but only one hour in the comparisons amongst teams (see below). We always applied time and not length restrictions to allow each team to search at its appropriate speed. We had two human teams, consisting of one expert and two students or one experienced person and three students and two otter scat detection dog teams, each consisting of one experienced handler, one student orienteer and one otter scat detection dog (Table 2). All students were introduced to otter scat detection on the first day.

Relative detection abilities

Each of the four teams visited the four previously cleared transects independently without exchange of people or knowledge. We used a rotation design, meaning that each team was once the first, second, third or fourth team on one transect. We limited the time spent on each transect to one hour so that each team could reach individual end points. Each team marked the start and end points to calculate transect lengths, as well as localities of otter scats on a plane-table sheet. For each scat found, the team described the site (on litter, on log, on rock, in bush, buried, den, latrine, other), scat condition (fresh, mostly fresh, dry, decayed), colour (yellow, green, black and all combinations thereof) and exposure (not exposed, exposed on natural structures, exposed on human structures) and estimated the scat size or gave the number of scats in the case of a hotspot. Each team also took two photos per scat, one close-up and one including the surroundings. These photos were used to assess whether the scats found by different teams were identical or different amongst search teams. During otter presence monitoring outside transects, protocolling was identical but no time and space restrictions applied. Dog teams additionally evaluated dog performance (Suppl. material 1: S1.4) and the likelihood of finding that scat without a dog (very low, moderate, high, uncertain).

We then analysed how many scats were found by each team per hour, the distance reached per hour and whether scat characteristics differed between scats found by human and dog teams using a t-test for count data and chi-test for categorical data. Data outside transects (presence monitoring) were only compared quantitatively since monitoring took place in different pond areas with different accessibilities and otter abundances.

Absolute detection abilities

The total number of otter scats was unknown throughout the field tests. Therefore, we estimated the total number of otter scats per transect using two independent visits of dog teams and the photos to assess which of the observed scats were "recaptures". Likewise, we estimated the total number of otter-like scats per transect, which comprises both otter scats and misidentified mink scats, using two independent visits of human teams. We calculated the minimum number of misidentified mink scats as the difference between the total number of otter and otter-like scats. Since we expect human teams to generally overlook very small or hidden scats (Long and MacKay 2012, Bonesi et al. 2013), the estimated total number of otter scats, based on human samples, would be smaller than that based on dog samples. Thus, if the estimated number of otter-like scats is larger than otter scats, the difference between both can only refer to a minimum number of misidentified mink scats.

We applied the Lincoln-Petersen estimator with the Chapman correction for small sample sizes (Seber 1982) using photos together with site descriptions for individual identification of scats. We shortened transects to the maximum common distance surveyed amongst all teams. We then calculated the detection rate for human and dog teams. All statistical analyses were conducted using R (R Core Team 2018) and the R-packages *exactRankTests* (Hothorn and Hornik 2013), *fifer* (Fife 2017) and *lme4* (Bates et al. 2015).

Results

Comparison of Eurasian otter and American mink scat

We found no combination of scat characteristics in wild samples of otter (n = 1273) and mink scats (n = 161) that would clearly indicate the species. Both species used hotspots and sometimes hotspots even contained scats of both species. However, otters used significantly more hotspots amongst all scats found (41%) than minks (23%) (χ^2 test, χ = 6.6, df = 1, p = 0.01). Likewise, otters used significantly more exposed surfaces (χ^2-test, χ = 45.22, df = 5, p < 0.001) despite all surfaces being used by both species. The amount of scat (χ^2-test, χ = 0.06, df = 2, p = 0.97) and the sliminess (χ^2 test, χ = 0.95, df = 2, p = 0.62) were not different between both species. We recognised the same amount of jelly scats ($\approx$ 20%) for both species, but while one-third of them were yellowish for both, the rest were greenish for minks and half greenish, half reddish for otters (χ^2-test, χ = 50.57, df = 2, p < 0.001).

Of the scat samples from captivity where both species fed on the same fish, we found that, of the otter scats, 18% were dark-green, 55% green, 18% yellow-green, and 9% yellow. Of the mink scats, 33% were black, 22% dark-green, 11% green, 22% yellow-green and 11% yellow. There was no statistical difference in scat colour between species (Fisher-test, df = 4, p = 0.11).

Species discrimination tests

We tested 3 beginners, 3 experienced persons and 4 experts. We found that the species assignment for the beginners were rather random with average accuracy of 0.58 (sensitivity from 0.45–0.64, specificity from 0.33–0.78) and average precision of 0.63. Experienced persons obtained an average accuracy of 0.72 (sensitivity from 0.63–0.81, specificity from 0.56–0.78) and an average precision of 0.74. Experts reached an average accuracy of 0.89 (sensitivity from 0.81–1, specificity from 0.67–1) and an average precision of 0.89 (Suppl. material 1: S1.5, Table S2). The overall average accuracy and precision were 0.75 and 0.77, respectively. The highest accuracy reached was 0.95.

None of the scat characteristics [species ($p = 0.43$), sex ($p = 0.06$), day sampled ($p = 0.08$), colour ($p = 0.89$), mass ($p = 0.41$)] clearly influenced correct assignment, with the almost significant sex-parameter being due to a male-biased sample and the almost significant day-sampled-parameter reflecting that the longer the species fed on fish only, the higher the chance of correct assignment. The colour was not significant because persons were assigning the species to different colours (Suppl. material 1: S1.3). However, the expertise of a test person clearly explained the assignment ($p = 0.004$), leading to an overall significant GLMM ($p = 0.04$). Correct assignment increased strongly with experience level.

In comparison, the four detection dogs showed average accuracy of 0.95 (from 0.9–1) with a sensitivity of 1 for all dogs and average specificity of 0.97 and 0.83 for the otter and mink scat detection dogs, respectively (Suppl. material 1: S1.5, Table S3). This resulted in average precision of 0.97 and 0.85 for the otter and mink scat detection dogs, respectively.

Relative detection abilities

Results per team and transect are given in Table 2. Otter detection dog teams found significantly more scats per hour (paired t-test, one-sided, $t = 4.25$, $df = 7$, $p = 0.002$, mean difference = 4.5 scats per hour) and were significantly faster (paired t-test, one-sided, $t = 2.33$, $df = 7$, $p = 0.02$, mean difference = 0.25 km/h) on standardised transects than human teams (Fig. 4). On average, dog teams found 12.5 and 13.5 scats per hour (paired t-test, two-sided, $t = -0.48$, $df = 3$, $p = 0.66$), but the team with the experienced dog Bagheera was slightly faster (1.1 km/h) than the team with the recruit Zammy (0.9 km/h) (paired t-test, two-sided, $t = 3.27$, $df = 3$, $p = 0.05$). The two human teams found eight and nine scats per hour on average (paired t-test, two-sided, $t = -0.41$, $df = 3$, $p = 0.71$) with an average speed of 0.6 and 0.7 km/h (paired t-test, two-sided, $t = -0.58$, $df = 3$, $p = 0.60$), respectively. During otter presence monitoring outside transects, dog teams were still faster (0.6 km/h vs. 0.4 km/h) and found more scats per hour (8.8 scats/h vs. 5.5 scats/h) (Table 3).

On standardised transects, human and detection dog teams found a total of 64 and 104 scats (Table 2). Of the scats found by each team, significantly more scats belonged to hotspots in the human (41%) than in the dog (25.7%) sample (χ^2-test, $\chi = 4.6$, $df = 1$,

Table 2. Results for transect comparisons amongst search teams. Time spent on each transect was limited to 1 h.

Team	Members	Transect	Distance reached [km]	Scats found	Scats per maximum common distance
Dogs1	Bagheera, AZ, 1 student	Biwatsch	1.12	12	5
		Mühlteich	0.41	15	9
		Entenschenke	0.98	13	13
		Langner	1.36	10	3
Dogs2	Zammy, AGS, 1 student	Biwatsch	1.06	10	5
		Mühlteich	0.36	22	12
		Entenschenke	0.80	13	12
		Langner	1.24	9	3
Humans1	Expert, 2 students	Biwatsch	1.06	3	2
		Mühlteich	0.20	10	10
		Entenschenke	0.68	13	13
		Langner	0.57	3*	0
Humans2	Experienced person, 3 students	Biwatsch	0.71	4	3
		Mühlteich	0.52	20	12
		Entenschenke	0.84	8	7
		Langner	0.80	3*	0

* Those three scats were found by both human teams at the beginning of the transect behind a thorny bush. The search of the dog teams started after that thorny bush so that the dogs could not pass them.

Table 3. Scat parameters evaluated for scats found by human teams and dog teams on standardised transects and during otter presence monitoring outside transects, with ranges given in brackets. Scat conditions do not add up to 100%, as on hotspots occasionally, several conditions could be found which are not included here.

Parameter	Human teams		Dog teams	
	transect	outside transect	transect	outside transect
Scats / hour	8.5 [3–20]	5.5 [1.3–10]	12.6 [9–22]	8.8 [5.2–12]
Km / hour	0.7 [0.2–1.1]	0.4 [0.2–0.6]	1 [0.4–1.4]	0.6 [0.5–0.7]
Scat condition	35% dry / decayed, 61% (mostly) fresh	73% dry / decayed, 22% (mostly) fresh	36% dry / decayed, 54% (mostly) fresh	48% dry / decayed, 42% (mostly) fresh
Scat size	61.8% large, 30.9% medium, 7.3% small	50.7% large, 34.2% medium, 15.1% small	32.6% large, 39.8% medium, 27.6% small	41% large, 34.6% medium, 24.4% small
Hotspots	41%	31.5%	25.7%	25.3%
Exposure	81.4%	79%	not evaluated	not evaluated

p = 0.03) and significantly more scats were large in the human (61.8% large, 30.9% medium, 7.3% small) than in the dog (32.6% large, 39.8% medium, 27.6% small) sample (χ^2-test, $\chi = 21.96$, df = 2, p < 0.001) (Table 3). During otter presence monitoring outside transects, scats found by human teams were preferably older and larger and belonged more often to hotspots. No such bias was detected for the scats found by dog teams, which found almost equal numbers of different sizes, ages and amount (Table 3). The likelihood of finding a specific scat without a detection dog was high in 22.5%, moderate in 31.7%, low in 39.4% and uncertain for 6.4% of the scats. Dogs showed high working abilities for their jobs and no change in their performance across time (Suppl. material 1: S1.4).
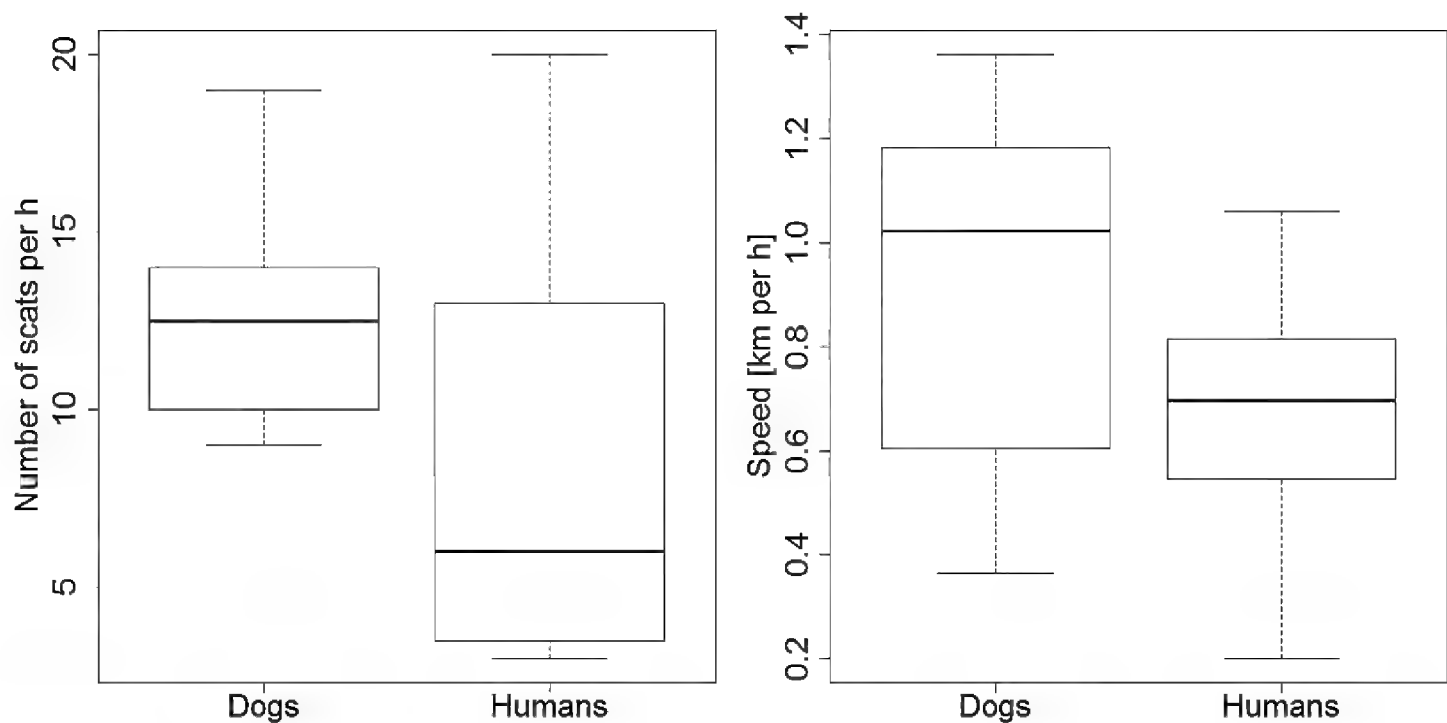
Figure 4. Comparison on the abilities of dog and human scat detection teams on standardised transects. The number of scats refers to the number of otter scats for detection dogs and otter-like scats for human teams.

Absolute detection abilities

Scat photo-identification performed well on one transect ("Mühlteich") with almost exclusively "naturally exposed" scats combined with scratch piles. Each team detected between 9 and 12 scats (Table 2) and the likelihood of finding a scat without a dog was almost always estimated as moderate to high. Average detection rate for detection dog and human teams was 0.60 [confidence interval (CI) 0.47–0.82] and 0.48 [CI 0.35–0.78], respectively. We estimated an abundance of 17.6 [CI 12.8–22.4] true otter scats and 22.8 [CI 14.1–31.6] otter-like scats. Thus, we assume that a minimum of 5 scats (22% [CI 9.2%-29.1%]), visually identified as otter, belonged to mink.

On two further transects, the relative detection rate for dog teams was 3–4 times higher than that of human teams (Table 2), making an estimation of scat abundances amongst all teams impossible. We therefore estimated scat detection rates within the maximum common distance for dog teams only and for human teams only. Detection rates for dog and human teams were estimated as 0.72 [CI 0.60–0.89] and 0.43 [CI 0.26–1] at "Langner" and 0.40 [CI 0.26–0.84] and 0.21 [CI 0.10–1] at "Biwatsch", respectively. Confidence intervals for the human samples should be treated with caution due to the very low sample sizes and extremely low "recapture" rates. On the fourth transect ("Entenschenke"), photo-identification failed due to otter scats being mostly on hotspots.

Discussion

In many parts of Europe, the Eurasian otter and American mink co-exist syntopically without out-competing each other (Harrington et al. 2009, Klenke et al. 2013, Lampa et al. 2015). When monitoring syntopic or sympatric species, scat differentiation is ex-

tremely valuable. As expected, we could not detect differences between characteristics of otter and mink scats that could be used for visual species identification. We detected more hotspots and more exposed scats for otters than for minks. However, Lampa et al. (2015) found fresh mink and otter scat on the same marking sites, making visual species assignment in the field impossible. Importantly, since mink scat was only sampled when identified as otter, results only represent the otter-like mink scats and might not reflect the overall use of hotspots and exposed areas in minks. Scats from minks that fed on other diet items, such as birds or small mammals, can easily be identified through remains of bones, hairs, feathers and a stronger smell and are thus not considered here. Rather, comparisons were intended to determine whether specific characteristics could be used for species identification when both species fed on identical diets, which is not the case. Without cross-checking species through genetic methods or detection dogs, a high number of otter-like scats of mink origin would significantly bias otter monitoring based on scat samples only (Lampa et al. 2015).

Often, scats of sympatric carnivores cannot be differentiated morphologically (MacKay et al. 2008). Our results indicated that human searchers showed the same false-positive detection rate in the field as in the species assignment test, which was comparable to an earlier study (Lampa et al. 2015). In other studies, even experts failed to reliably distinguish pine marten (*Martes martes*) scats from those of foxes (*Vulpes vulpes*) (Davison et al. 2002) or at detecting mink scat amongst other mustelid species (Harrington et al. 2010). Our species discrimination tests showed that beginners assigned the species randomly, while experts showed high accuracy nearly equalling the accuracy which dogs already showed during the discrimination training (Suppl. material 1: S1.2). Moreover, it took years to become an expert (given that the monitoring time per year is limited), while the training of detection dogs just needed a few months and was, thus, a much faster method. After completing discrimination training, detection dogs clearly outperformed the abilities of any expert. However, we suggest that dogs performing only one search task (e.g. wildlife detection) might reach higher accuracy faster, since each search task (e.g. being a search and rescue dog in parallel, as for our mink detection dogs) would require a huge amount of training simultaneously which is unlikely to be adequately addressed.

In comparison, Oldenburg et al. (2016) also tested specificity and sensitivity of one otter scat detection dog, but they used unrelated decoy species with entirely different diets. In their final test, the dog showed a sensitivity of 1, as ours did, but it also alerted at rabbits and hares with a specificity of 0.95, summing to an accuracy of 0.96 and a precision of 0.7. In other studies, scat detection dogs showed a specificity of 0.93 for fisher (*Martes pennanti*) (Long et al. 2007a) and 1 for kit fox (*Vulpes macrotis*) (Smith et al. 2001), stressing the high discrimination abilities of dogs. Hence, using detection dogs for species identity will allow subsequent laboratory analyses to be species-specific and avoid spending time and money on laboratory work of the wrong species.

Furthermore, our detection dog teams significantly increased the efficiency of field monitoring by detecting more scats and reaching a longer distance than human teams per hour. Moreover, detection dogs were twice as fast as human teams despite their handlers having to stop three to four times more often to protocol samples and fill out

a longer protocol. In North America, Richards (2016) also compared the abilities of detection dogs with humans, both searching for the scat of the American mink and North American river otter (*Lontra canadensis*) simultaneously. On a 2.4 km shoreline transect searched for 2 h, an expert and a detection dog found one and eleven scats, respectively and, on another transect of 5.5 km length, a biologist could not detect any scat while detection dogs found 16. Likewise, Smith et al. (2001) reported that a detection dog found up to four times more kit fox scats than an experienced searcher. In addition, our dogs were less selective in the area they searched, as they detected scats on logs reaching into the water or below leaves and in grass impossible to detect visually. Finally, while humans detected scats with characteristics that increased their visibility (larger, more exposed, older and thus brighter), scats found by dogs were much less skewed towards a particular size or age. Scat sampling, biased to more exposed or larger scats, could bias ecological analyses (Birks et al. 2005, Bonesi et al. 2013). Thus, their ability to find many scats over large areas, including small and hidden ones and their lesser spatial and sampling bias, make detection dogs particularly helpful for studying marten species (Long and MacKay 2012). Typical problems when monitoring martens through scat monitoring, including biased searching ("knowing where to search") and scat differentiation (Birks et al. 2005), can be overcome by using detection dogs.

In this study, different detection rates of the detection dogs were likely due to different vegetation conditions amongst transects, with the lowest detection rate at the transect with many bushes. Nevertheless, detection rates of detection dogs were always twice as high as visual detection rates. In studies with other mammals, scat detection dogs also outperformed other monitoring methods, such as cameras, hair snares and scent stations (Harrison 2006, Long et al. 2007b, Tom 2012) and required much less time to ascertain species presence (Long et al. 2007a, Clare et al. 2015). Importantly, the frequency of detection and the accuracy can also vary amongst dogs, dog-handler-teams, day, weather and place (MacKay et al. 2008), despite variation can be limited with proper adjustments (Leigh and Dominick 2015). It is therefore necessary to adapt the training specifically to the dog and the given field conditions (Woollett (Smith) et al. 2014).

Conclusion

Using species-specific scat detection dogs is saving time during species monitoring and allows subsequent laboratory analyses to be species-specific. Even for related species feeding on identical diets, scat detection dogs showed highly accurate and precise species discrimination abilities, which outperformed those of experts. Detection dogs were also reliable in the field and more efficient than humans. Moreover, detection dogs provided more precise and unbiased information about the target species than visual searching through, for example, detecting hidden or very small samples. With proper training provided, their use is highly recommended for species monitoring through scats and will significantly contribute to optimal conservation of threatened and management of conflict species.

Acknowledgements

References

Arrendal J, Vilà C, Björklund M (2007) Reliability of noninvasive genetic census of otters compared to field censuses. Conservation Genetics 8(5): 1097–1107. https://doi.org/10.1007/s10592-006-9266-y

Bates D, Maechler M, Bolker B, Walker S (2015) Fitting linear mixed-effects models using lme4. Journal of Statistical Software 67(1): 1–48. https://doi.org/10.18637/jss.v067.i01

Birks J, Messenger J, Braithwaite T, Davison A, Brookes R, Strachan C (2005) Are scat surveys a reliable method for assessing distributon and population status of pine martens? In: Harrison DJ, Fuller A, Proulx G (Eds) Martens and Fishers (Martes) in Human-altered Environments. Springer, Boston, 235–252. https://doi.org/10.1007/0-387-22691-5_12

Bonesi L, Hale M, Macdonald DW (2013) Lessons from the use of non-invasive genetic sampling as a way to estimate Eurasian otter population size and sex ratio. Acta Theriologica 58(2): 157–168. https://doi.org/10.1007/s13364-012-0118-5

Clare JDJ, Anderson EM, MacFarland DM, Sloss BL (2015) Comparing the costs and detectability of bobcat using scat-detecting dog and remote camera surveys in central Wisconsin. Wildlife Society Bulletin 39(1): 210–217. https://doi.org/10.1002/wsb.502

Dahlgren DK, Elmore RD, Smith DA, Hurt A, Arnett EB, Connelly JW (2012) Use of dogs in wildlife research and management. In: Silvy NJ (Ed.) The wildlife techniques manual. Seventh edition The John Hopkins University Press, Baltimore, 140–153.

Davison A, Birks JDS, Brookes RC, Braithwaite TC, Messenger JE (2002) On the origin of faeces: Morphological versus molecular methods for surveying rare carnivores from their scats. Journal of Zoology 257(2): 141–143. https://doi.org/10.1017/S0952836902000730

DeMatteo KE, Davenport B, Wilson LE (2019) Back to the basics with conservation detection dogs: fundamentals for success. Wildlife Biology 2019: wlb.00584. https://doi.org/10.2981/wlb.00584

Fife D (2017) fifer: A biostatisticians toolbox for various activities, including plotting, data cleanup, and data analysis. R package version 1.1. https://cran.r-project.org/package=fifer

Gadbois S, Reeve C (2014) Canine olfaction: scent, sign, and situation. In: Horowitz A (Ed.) Domestic dog cognition and behavior: the scientific study of Canis familiaris. Springer, Berlin-Heidelberg, 3–29. https://doi.org/10.1007/978-3-642-53994-7_1

Geidezis L (1996) Food availability versus food utilization by otters (*Lutra lutra* L.) in the Oberlausitz ponland in Saxony, Eastern Germany. IUCN Otter Specialist Group Bulletin 13: 58–67.

Hájková P, Zemanová B, Roche K, Hájek B (2009) An evaluation of field and noninvasive genetic methods for estimating Eurasian otter population size. Conservation Genetics 10(6): 1667–1681. https://doi.org/10.1007/s10592-008-9745-4

Harrington LA, Harrington AL, Yamaguchi N, Thom MD, Ferreras P, Windham TR, Macdonald DW (2009) The impact of native competitors on an alien invasive: Temporal niche shifts to avoid interspecific aggression? Ecology 90(5): 1207–1216. https://doi.org/10.1890/08-0302.1

Harrington LA, Harrington AL, Hughes J, Stirling D, Macdonald DW (2010) The accuracy of scat identification in distribution surveys: American mink, *Neovison vison*, in the northern highlands of Scotland. European Journal of Wildlife Research 56(3): 377–384. https://doi.org/10.1007/s10344-009-0328-6

Harrison RL (2006) A comparison of survey methods for detecting bobcats. Wildlife Society Bulletin 34(2): 548–552. https://doi.org/10.2193/0091-7648(2006)34[548:ACOSMF]2.0.CO;2

Hauer S, Ansorge H, Zöphel U (2009) Atlas der Säugetiere Sachsens. Sächsisches Landesamt für Umwelt, Landwirtschaft und Geologie, Dresden, Germany.

Henle K, Ring I, Klenke RA, Kranz A, Jepsen N, Rauschmayer F (2013) Introduction. In: Klenke RA, Ring I, Kranz A, Jepsen N, Rauschmayer F, Henle K (Eds) Human-Wildlife Conflicts in Europe. Springer, Berlin-Heidelberg, 5–11. https://doi.org/10.1007/978-3-540-34789-7_2

Holm S (1979) A simple sequentially rejective multiple test procedure. Scandinavian Journal of Statistics 6: 65–70.

Hothorn T, Hornik K (2013) exactRankTests: Exact distributions for rank and permutation tests. R package version 0.8-27. https://cran.r-project.org/package=exactRankTests

Hurt A, Davenport M-B, Greene E (2000) Training dogs to distinguish between black bear (Ursus americanus) and grizzly bear (Ursus arctos) feces. Under-Graduate Biology Journal. University of Montana, USA.

Jedrzejewska B, Sidorovich VE, Pikulik MM, Jedrzejewski W (2001) Feeding habits of the otter and the American mink in Białowieża Primeval Forest (Poland) compared to other Eurasian populations. Ecography 24(2): 165–180. https://doi.org/10.1034/j.1600-0587.2001.240207.x

Kerley LL (2004) Scent dog monitoring of Amur tiger (2003-0087-018) March 1, 2003 – March 1, 2004. Lazovsky State Nature Zapovednlk. Lazo, Primorski Region, Russia.

Klenke RA, Ring I, Schwerdtner Mánez K, Habighorst R, Weiss V, Wittmer H, Gruber B, Lampa S, Henle K (2013) Otters in Saxony: A story of successful conflict resolution. In: Klenke RA, Ring I, Kranz A, Jepsen N, Rauschmayer F, Henle K (Eds) Human-Wildlife Conflicts in Europe. Springer, Berlin-Heidelberg, 107–140. https://doi.org/10.1007/978-3-540-34789-7_6

Kranz A (2000) Otters (*Lutra lutra*) increasing in Central Europe: From the threat of extinction to locally perceived overpopulation? Mammalia 64(4): 357–368. https://doi.org/10.1515/mamm.2000.64.4.357

Kruuk H (2006) Otters – ecology, behaviour and conservation. Oxford University Press, Oxford, UK, New York, USA. https://doi.org/10.1093/acprof:oso/9780198565871.003.0014

Lampa S, Mihoub JB, Gruber B, Klenke R, Henle K (2015) Non-invasive genetic mark-recapture as a means to study population sizes and marking behaviour of the elusive Eurasian otter (*Lutra lutra*). PLoS One 10(5): e0125684. https://doi.org/10.1371/journal.pone.0125684

Leigh KA, Dominick M (2015) An assessment of the effects of habitat structure on the scat finding performance of a wildlife detection dog. Methods in Ecology and Evolution 6(7): 745–752. https://doi.org/10.1111/2041-210X.12374

Long RA, MacKay P (2012) Noninvasive methods for surveying martens, sables, and fishers. In: Aubry KB, Zielinski WJ, Raphael MG, Proulx G, Buskirk SW (Eds) Biology and conservation of martens, sables, and fishers. A new synthesis. Cornell University Press, New York, 320–342. https://doi.org/10.7591/9780801466076-017

Long RA, Donovan TM, MacKay P, Zielinski WJ, Buzas JS (2007a) Effectiveness of scat detection dogs for detecting forest carnivores. The Journal of Wildlife Management 71(6): 2007–2017. https://doi.org/10.2193/2006-230

Long RA, Donovan TM, MacKay P, Zielinski WJ, Buzas JS (2007b) Comparing scat detection dogs, cameras, and hair snares for surveying carnivores. The Journal of Wildlife Management 71(6): 2018–2025. https://doi.org/10.2193/2006-292

MacKay P, Smith DA, Long RA, Parker M (2008) Scat detection dogs. In: Long RA, MacKay P, Zielinski WJ, Ray JC (Eds) Noninvasive survey methods for carnivores. Island Press, Washington-Covelo-London, 183–222.

Mason CF, Macdonald SM (1987) The use of spraints for surveying otter Lutra lutra populations: An evaluation. Biological Conservation 41(3): 167–177. https://doi.org/10.1016/0006-3207(87)90100-5

McKay S (2014) Scent-matching of olfactory samples via proxy by domestic dogs Canis lupus familiaris. University of the Witwatersrand, Johannesburg, South Africa.

Oldenburg Jr C, Schoon A, Heitkönig IMA (2016) Wildlife detection dog training: A case study on achieving generalization between target odor variations while retaining specificity. Journal of Veterinary Behavior 13: 34–38. https://doi.org/10.1016/j.jveb.2016.03.008

Poledníková K, Kranz A, Poldník L, Mysiak J (2013) Otters causing conflicts. In: Klenke RA, Ring I, Kranz A, Jepsen N, Rauschmayer F, Henle K (Eds) Human-Wildlife Conflicts in Europe. Springer, Berlin-Heidelberg, 81–106. https://doi.org/10.1007/978-3-540-34789-7_5

R Core Team (2018) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna. http://www.r-project.org/

Richards NL (2016) Using conservation detection dogs to monitor environmental contaminants in aquatic sentinel species and freshwater ecosystems. Working Dogs for Conservation.

Seber GA (1982) The estimation of animal abundance and related parameters. Griffin, London.

Smith DA, Ralls K, Davenport B, Adams B, Maldonado JE (2001) Canine Assistants for Conservationists. Science 291(5503): 435. https://doi.org/10.1126/science.291.5503.435B

Smith DA, Ralls KAH, Adams B, Parker M, Davenport B, Smith M, Maldonado J (2003) Detection and accuracy rates of dogs trained to find scats of San Joaquin kit foxes (*Vulpes macrotis mutica*). Animal Conservation 6(4): 339–346. https://doi.org/10.1017/S136794300300341X

Tom BM (2012) A comparison of noninvasive survey methods for monitoring mesocarnivore populations in Kentucky. Theses and Dissertations – Forestry Paper 10: 1–140.

Wasser SK, Davenport B, Ramage ER, Hunt KE, Parker M, Clarke C, Stenhouse G (2004) Scat detection dogs in wildlife research and management: Application to grizzly and black bears in the Yellowhead Ecosystem, Alberta, Canada. Canadian Journal of Zoology 82(3): 475–492. https://doi.org/10.1139/z04-020

Woollett SDA, Hurt A, Richards NL (2014) The current and future roles of free-ranging detection dogs in conservation efforts. In: Gompper ME (Ed.) Free-ranging dogs & wildlife conservation. Oxford University Press, Oxford, 239–264. https://doi.org/10.1093/acprof:osobl/9780199663217.003.0010

Supplementary material I

Appendix S1. Additional information and analyses

Authors: Annegret Grimm-Seyfarth, Aleksandra Zarzycka, Teresa Nitz, Lisa Heynig, Nadine Weissheimer, Simone Lampa, Reinhard Klenke

Data type: statistical data

Explanation note: **S1.1** Original Eurasian otter monitoring study; **S1.2** Detection dogs pre-test and discrimination training; **S1.3** Do humans assign species based on color? **S1.4** Evaluation of dog performance in the field; **S1.5** Species assignments for humans and dogs.

Copyright notice: This dataset is made available under the Open Database License (http://opendatacommons.org/licenses/odbl/1.0/). The Open Database License (ODbL) is a license agreement intended to allow users to freely share, modify, and use this Dataset while maintaining this same freedom for others, provided that the original source and author(s) are credited.

Link: https://doi.org/10.3897/natureconservation.37.48208.suppl1